

# Ölçek Geliştirme Çalışmalarında Kullanılan İçerik Geçerliliğine Genel Bir Bakış

## An Overview Evaluation of the Content Validity Used in Scale Development Studies

• İsmet DOĞAN<sup>a</sup>,  
• Nurhan DOĞAN<sup>a</sup>

<sup>a</sup> Biyoistatistik ve Tıbbi Bilişim AD,  
Afyonkarahisar Sağlık Bilimleri Üniversitesi  
Tıp Fakültesi,  
Afyonkarahisar, TÜRKİYE

Received: 07.03.2019  
Received in revised form: 08.04.2019  
Accepted: 09.04.2019  
Available Online: 11.04.2019

Correspondence:  
Nurhan DOĞAN  
Afyonkarahisar Sağlık Bilimleri Üniversitesi  
Tıp Fakültesi,  
Biyoistatistik ve Tıbbi Bilişim AD,  
Afyonkarahisar,  
TÜRKİYE /TURKEY  
nurhandogan@hotmail.com

**ÖZET Amaç:** Makalenin amacı, içerik geçerliliğinde kullanılan nicel ölçütleri tanıtmak ve Lawshe'nin içerik geçerlilik oranı için asgari uzman sayısını belirlemektir. Çalışmada içerik geçerliliğinin tanımı, önemi, kavramsal temeli ve fonksiyonel doğası incelenmiştir. **Gereç ve Yöntemler:** Bir testin (alet, ölçek ya da anketin) kabul edilebilir olup olmadığını söyleyebilmek için göz önünde bulundurulmuş beş kriter güvenilirlik, geçerlik, objektiflik, pratiklik ve ekonomiklik olarak sıralanmaktadır. Kabul edilebilir bir testin geliştirilmesinde geçerlilik ve güvenilirlik özellikle üzerinde durulan iki önemli kavramdır ve ölçüm geçerliliği yaklaşık 100 yıldır bilim adamlarının dikkatini çeken önemli ve karmaşık bir konudur. Geçerlilik, araştırmacıların elde ettiği yanıtlardan elde edilen çıkarımların bir özelliğidir ve bu özellik araştırmacıların bu yanıtları elde etmek için kullandıkları araçlara ait olmayıp sabit bir özelliktir. **Bulgular:** Gerek yayınlanmamış ve piyasada mevcut ölçeklerin gerekse diğer psikometrik ölçeklerin kullanım sıklığında son yıllarda artış görülmektedir. Ne yazık ki, bu ölçeklerin birçoğunun geçerliliği ve güvenilirliği ile ilgili bilgiler ya elde edilememektedir, ya yetersizdir ya da uygun değildir. Psikometrik ölçekte içerik geçerliliğinin önemi ve içerik geçerliliğinin güvenilirlik ile ilişkisi, ölçek geliştirmede içerik geçerliliğini önemli bir adım haline getirmiştir. **Sonuç:** İçerik geçerliliği yargıların ölçülmesi için geliştirilen alternatif yöntemlerle mukayese edildiğinde, Lawshe yöntemi yalnızca basit hesaplamalar gerektiren ve kritik bir eşik değeri belirlemek için kullanıcılara kullanımı kolay bir tablo sunan basit ve kullanıcı dostu bir yöntemdir.

**Anahtar Kelimeler:** İçerik geçerliliği; içerik geçerlilik oranı; içerik geçerlilik indeksi; test oluşturma; ölçek geliştirme

**ABSTRACT Objective:** The purpose of the article is to introduce quantitative criteria used for content validity and to determine the minimum number of experts for content validity of Lawshe. The definition, importance, conceptual basis, and functional nature of content validity were examined in the article. **Material and Methods:** There are five criteria that are used to know whether or not the test (instrument, scale or questionnaire) can be considered qualified namely reliability, validity, objectivity, practicality, and economy. Validity and reliability are two important things related to constructing such kind of test. Validity of measurement is an important and complex issue that has received much attention from scientists for nearly 100 years. Validity is a property of the inferences investigators make from the responses they obtain and not a fixed property of the instruments used to obtain those responses. **Results:** Both unpublished and commercially available scales and other psychometric devices yielding ordinal level measurements appear to have increased in frequency of usage during recent years. Unfortunately, information concerning the reliability and validity of many of these instruments is either unavailable, insufficient, or inappropriate. The importance of content validity in the instrument psychometric and its relevance with reliability, have made it an essential step in the instrument development. **Conclusion:** Compared with alternative methods for quantifying content validity judgements, the Lawshe method is straightforward and user-friendly, requiring only simple computations and providing a table for determining a critical cutoff value.

**Keywords:** Content validity; content validity ratio; content validity index; test construction; instrument development

**B**ilimsel çalışmaların çoğunda geçerli ve güvenilir ölçeklerin gerekli olduğu karmaşık yapılar incelendiğinden, ölçek geliştirme ile uğraşan araştırmacıların ölçeğin güvenilirliği ve geçerliliği hakkında kapsamlı bilgi sunmaları beklenmektedir. Bir araştırmanın sonuçlarına güven duymak için kullanılan ölçeğin ölçmek istediği şeyi tutarlı bir şekilde ölçmesi sağlanmalıdır.<sup>1</sup> Somut ve soyut özellikleri ölçmek için kullanılacak ölçek, standart bir ölçme aracı olmalıdır. Bir ölçeğin standart ölçek olabilmesi için ise güvenilirlik ve geçerlilik gibi iki temel özelliğe ve uygun bilgiyi üretme yeteneğine sahip olması gerekir.<sup>2</sup> Güvenilirlik, ölçüm ile elde edilen sonuçların tekrarlanabilirlik derecesini belirtmektedir ve ölçeğin geçerliliğine önemli ölçüde katkıda bulunmasına rağmen ölçek geçerliliği için yeterli koşul değildir. Güvenilirlik eksikliği, gözlemciler veya ölçüm araçları arasındaki farklılıktan veya ölçülen öznenin kararsızlığından kaynaklanabilir. Güvenilirliğin eşdeğerlik, istikrar ve iç tutarlılık (homojenlik) gibi üç yönü söz konusudur.<sup>3</sup> Geçerlilik ise bir ölçeğin incelenen fenomenin özelliklerini ölçme kabiliyeti olarak tanımlanmaktadır ve ölçeğin tercih edilmesinde veya kullanılmasında hayati bir faktördür. Geçerlilik içerik, yapı ve kritere bağlı geçerlilik olmak üzere üç genel forma sahiptir. İçerik geçerliliği, diğer geçerlilikler için bir ön koşul olduğundan ölçek geliştirmede en yüksek önceliği alması gerekir. Geçerlilik, ölçekten ziyade özel bir katılımcı grubu üzerinde belirli bir amaç için kullanılan ölçekten elde edilen puanlar ile ilgili olduğundan, her çalışma için geçerlilik kanıtı elde edilmelidir. Bir ölçek içerik geçerliliğine sahip değilse, bu ölçek için güvenilirlik sağlamak mümkün değildir.<sup>4</sup> Ne istatistiksel zerafet, ne de sofistike ölçüm stratejileri, geçerlilik yerine geçemez. İçerik geçerliliğinin değerlendirilmesinde sistematik ve derinlemesine yaklaşımlar gereklidir.<sup>5</sup> İçerik geçerliliği bir ölçekteki maddelerin ne kadar iyi örneklendiğini, kritere bağlı geçerlilik kriteri ile ölçüm arasındaki ilişkiyi, yapı geçerliliği ise ölçeğin teorik bir yapı veya özelliği ne kadar ölçtüğünü belirtmektedir.<sup>6</sup> İçerik geçerliliği öznel yargılara dayandığından, bu tür geçerliliğin belirlenmesi diğer geçerlilik türlerine göre daha öznedir.<sup>7</sup> Her ne kadar yeni bir ölçeğin kritere bağlı geçerliliği ile yapı geçerliliği özellikle önemli olarak kabul edilse de, ölçeğin kalitesiyle ilgili sonuçların çıkarılmasında içerik geçerliliği ile ilgili bilgiler de gerekli görülmektedir.<sup>8</sup> Bir ölçeğin geçerliliği tartışılırken öncelikle, “ölçek ölçülmek istenen fenomeni ölçebiliyor mu?” sorusunun cevaplanması gerekir. Herhangi bir ölçeğin geçerliliğini tespit etmenin ilk aşaması, içerik geçerliliği değerlendirmesiyle başlar.<sup>9</sup> İçerik geçerliliği, bir ölçeğin tasarımı sırasında sağlanması gereken ilk geçerlilik türüdür ve diğer geçerlilikler için ön koşul olması ve ölçeğin güvenilirliğiyle olan yakın ilişkisinden dolayı ölçek tasarımında önemli bir yere sahiptir.<sup>10</sup> İçerik geçerliliği, ölçek geliştirmede önemli bir adımdır, çünkü soyut kavramları gözlemlenebilir ve ölçülebilir göstergelerle ilişkilendirmeye yönelik başlangıç mekanizmasıdır.<sup>11</sup> Sireci (1998), içerik geçerliliğini alan tanımı, alan temsili, alan uygunluğu ve test geliştirme sürecinin uygunluğu biçimindeki ölçek kalitesini belirleyen dört unsuru da kapsayacak bir tanım yapmıştır. Belirli bir amaç için bir ölçeğin kullanımı ile ilgili yapılacak değerlendirmede bu dört unsurun önemli olduğu konusunda fikir birliği olsa da, geçerlilik ile ilgilenen birçok teorisyen içerik geçerliliğinin teknik olarak doğru bir terim olmadığını, çünkü geçerliliğin ölçek puanlarının yorumlanmasına atıfta bulunduğunu iddia etmektedir.<sup>12</sup> Ölçek geliştirme ile ilgili standartların belirlenmesi için yapılan ilk çalışmadan itibaren, geliştirilen ölçeklerin değerlendirilen fenomenin içerik alanını yeterince temsil ettiğine dair kanıt sağlama ihtiyacı her geçen gün artmaktadır. Ölçek geliştirme ile ilgili standartlar, ölçek geliştirmede oldukça önemli yer tutan ve tanım, temsil edilebilirlik ve alaka düzeyi şeklinde ifade edilen üç kavrama verilen artan önemi yansıtmaktadır. Bu kavramlar, içerik geçerliliğinin değerlendirilmesi için değişmeden kalan temel kavramsal yapıyı oluşturmaktadır.<sup>13</sup> İçerik geçerliliği, bir ölçeğin ölçmesi beklenen içerik alanını ölçtüğü güvencesini sağlamayı amaçlayan bir süreçtir.<sup>14</sup> Ölçekte yer alacak maddelerin tümü değerlendirme aracının amaçlanan içeriğini yansıtmalıdır. Dolayısıyla araştırmacılar içerik geçerlilik çalışması yaparak çok değerli bilgiler elde edebilirler. Bir uzmanlar panelinin kullanılması, yeni geliştirilen ölçeğin kalitesi ve ölçekte yer alacak her bir madde ile ilgili yapıcı geri bildirim sağlar.<sup>6</sup> Bir ölçeğin içerik geçerliliğini değerlendirmek, ölçeğin yapı ge-

çerliliğini arttırmada kritik bir ön adımdır. Bu nedenle içerik geçerliliği, yüksek kaliteli ölçümler isteyen klinisyenler ve araştırmacılar için önemli bir konudur. İçerik geçerliliği, yapı geçerliliğinin sağlanmasına yardımcı olur ve araştırmacılara ölçekler hakkında güven verir. Yeni bir ölçek geliştirecek araştırmacılardan, ölçeğinin ve ölçekte yer alacak maddelerin geçerli olduğuna dair kanıt sağlamaları beklenmektedir.<sup>15,16</sup> İçerik geçerliliği için literatürde farklı tanımlar olmasına rağmen, birçoğu bunu bir değerlendirme aracının unsurlarının ilgililenen yapıyı temsil etme derecesi olarak tanımlamaktadır.<sup>17</sup>

Ululararası literatürde özellikle sağlık alanında çok sayıda ölçek kullanılmaktadır. Türkiye Ölçme Araçları Dizini'nde Tıp Bilimleri başlığı altında 23, Temel Tıp Bilimleri başlığı altında 107, Dahili Tıp Bilimleri başlığı altında 204 ve Cerrahi Tıp Bilimleri başlığı altında ise 97 adet ölçek bulunmaktadır. Ölçek geliştirmede, içerik geçerliliği kritik bir adım ve soyut kavramları görünür ve ölçülebilir indekslere bağlayan bir tetikleyici mekanizma olmasına rağmen, yüzeysel olarak çalışılmaktadır. Bu sorun, tıbbi araştırma literatüründeki içerik geçerliliğini değerlendirmek için kullanılan yöntemlerin derinden yönlendirilmesinden kaynaklanmaktadır. Bu çalışmada içerik geçerliliğinin tanımı, önemi, kavramsal temeli ve işlevsel niteliği incelenmektedir. Makalenin amacı, içerik geçerliliğinde kullanılan nicel ölçütleri tanıtmak ve Lawshe'nin içerik geçerlilik oranı için asgari uzman sayısını belirlemektir.

## GEREÇ VE YÖNTEMLER

Ölçek geliştirilmesi ve değerlendirilmesi karmaşık ve uzun bir süreçtir. Bir ölçeğin içerik geçerliliği ne kadar yüksek olursa, ölçülmesi hedeflenen fenomenin ölçümü o kadar doğru olur.<sup>18</sup> İçerik geçerliliğini sağlamanın en yaygın yolu, bir ölçekte yer alan maddelerin önemini belirleyen bir konu uzmanı paneli oluşturmaktır. İçerik geçerliliği için uzmanlar tarafından ölçekte yer alması planlanan maddelerin incelenmesinden elde edilen nicel ve nitel göstergeler, ölçek geliştirme aşamasında yanlış adımların ve düzeltilen içeriğin belirlenmesinde faydalı olabilir.<sup>19</sup> İçerik geçerliliği tahmin edilirken nicel bir ölçüt kullanılması esastır. Uzmanlar tarafından içerik geçerliliğinde kullanılan bu ölçütler içerik geçerlilik indeksi (Content Validity Index) ve içerik geçerlilik oranı (Content Validity Ratio)'dır. İçerik geçerlilik oranı (İGO), ayrı ayrı her bir maddenin ölçekte yer alıp almayacağına karar verilmesinde uluslararası kabul gören bir ölçüttür. İçerik geçerlilik indeksi (İGİ) ise, nihai ölçekte yer alan tüm maddeler için ortalama İGO'dur.<sup>20</sup> Bir başka ifade ile İGO her bir maddenin gerekli olup olmadığını belirlemek için İGİ ise ölçekte yer alan her bir maddenin kullanılan ölçekle olan ilişkisini belirlemek için kullanılır.<sup>21</sup> İGİ uzmanların maddelerin uygunluğu ve açıklığı konusunda hemfikir olma derecelerinden yararlanılarak hesaplanmaktadır.<sup>1</sup> Her bir maddenin alaka düzeyinin dört puanlı sıralı bir derecelendirme (1 = alakalı değil, 2 = biraz alakalı, 3 = oldukça alakalı, 4 = yüksek düzeyde alakalı) ile değerlendirildiği madde içerik geçerlilik indeksi (M-İGİ), Lynn tarafından geliştirilmiştir.<sup>22</sup> Polit ve Beck tarafından önerilen ve ölçek içerik geçerlilik indeksi (Ö-İGİ) olarak isimlendirilen yöntemde ise derecelendirmeler, 1 = alakasız, 2 = biraz alakalı, 3 = alakalı ve 4 = son derece alakalı şeklindedir. Açıkça anlaşılacağı üzere 1 puan maddenin önemsiz olduğunu, dört puan ise maddenin aşırı derecede alakalı olduğunu belirtir. Gerek Lynn gerekse Polit ve Beck tarafından tanımlanan indekslerin her ikisinde de her bir maddenin uygunluğu için, 3 veya 4 puan veren uzman sayısı toplam uzman sayısına bölünür. Polit ve Beck tarafından içerik geçerlilik verilerinin yorumlanabilirliğini artırmak için Ö-İGİ'nin de hesaplanması gerektiği önerilmiştir. Ö-İGİ istatistiği ortalama madde kalitesini temsil etmekte ve uzmanlar arasındaki farklı fikirleri açıklamaktadır. M-İGİ'ler ile karşılaştırıldığında Ö-İGİ'ler, uzmanlar tarafından bir ölçekte 3 veya 4 puan verilen tüm maddelerin oranını göstermektedir.<sup>23</sup> Ö-İGİ ve M-İGİ'ler ile ilgili örnek bir hesaplama Tablo 1'de gösterilmiştir. Tablo 1, 10 maddelik bir ölçek için altı uzmana ait alaka derecelerini göstermektedir. Bu örnekte, altı uzmanın hepsi ilgili 10 maddenin tamamını derecelendirmiştir. 10 maddeden sadece 4'ü tüm uzmanlar tarafından 3 veya 4 alaka derecesi almıştır. (Tablo 1)

**TABLO 1:** Madde içerik geçerlilik indeksi ve ölçek içerik geçerlilik indeksi için örnek hesaplamalar.

Madde	Uzman 1	Uzman 2	Uzman 3	Uzman 4	Uzman 5	Uzman 6	Hemfikir Olan Uzman Sayısı	M-İGİ
1	X	X	X	X	X	X	6	1.00
2	X	-	X	X	-	X	4	0.67
3	X	X	-	X	X	X	5	0.83
4	X	X	X	X	X	X	6	1.00
5	-	X	X	X	-	-	3	0.50
6	X	X	X	X	X	-	5	0.83
7	X	X	X	X	X	X	6	1.00
8	X	X	X	-	X	X	5	0.83
9	X	X	X	X	X	X	6	1.00
10	-	X	-	-	-	X	2	0.33
Alakalılık Oranı	0.80	0.90	0.80	0.80	0.70	0.80	Ö-İGİ = 0.40	

M-İGİ: Madde içerik geçerlilik indeksi, Ö-İGİ: Ölçek içerik geçerlilik indeksi

Ö-İGİ değeri, panelde yer alan uzman sayısına karşı hassastır. Panelde yer alan uzman sayısı arttıkça düşük bir Ö-İGİ olma olasılığı o kadar artar. Eğer panelde yer alan uzman sayısı 5'ten fazla ise, M-İGİ değeri 0,78'den düşük olmamalıdır.<sup>24</sup>

İGO, içerik geçerliliğini ölçmek için kullanılan en eski ve en yaygın yöntemlerden biridir. Lawshe (1975) tarafından önerilen İGO değerinin hesaplanabilmesi için, ölçekte yer alan her bir maddenin “gerekli”, “faydalı, ancak gerekli değil” ve “gerekli değil” biçiminde derecelendirilmesi gerekir. Panelde yer alan uzmanlardan elde edilen bilgiler kullanılarak İGO değeri;

$$iGO = \frac{n_e - (N/2)}{N/2} \quad (1)$$

$n_e$ : "gerekli" cevabını veren uzman sayısı,

$N$ : panelde yer alan toplam uzman sayısı,

eşitliği kullanılarak hesaplanır. İGO değerleri -1 (mükemmel uyumsuzluk) ile +1 (mükemmel uyuma) arasında değişmektedir. Panelde yer alan uzmanların tümü “gerekli” cevabını verdiğinde İGO'nun değeri 1, “gerekli” cevabını veren uzman sayısı yarıdan az olduğunda negatif ve “gerekli” cevabını veren uzman sayısı yarıdan fazla ise  $0 \leq iGO \leq 1$  olur.<sup>14,25</sup> Elde edilen İGO değerine göre,

■ Herhangi bir madde için panelde yer alan uzmanların tamamı “gerekli değil” cevabını vermiş ise o maddenin tamamen gereksiz olduğu,

■ Herhangi bir madde için panelde yer alan uzmanların tamamı “faydalı ancak gerekli değil” cevabını vermiş ise o maddenin önemli ölçüde gerekli olduğu,

■ Herhangi bir madde için “gerekli” cevabını veren uzman sayısı yarıdan fazla ise o maddenin belirli bir geçerlilik değerine sahip olduğu, “gerekli” cevabını veren uzman sayısı arttıkça maddenin geçerlilik değerinin de yükseleceği,

yorumları yapılabilir.<sup>26</sup> Lawshe (1975) tarafından, panelist sayısına göre olması gereken en düşük İGO değerleri Tablo 2'de sunulmuştur.<sup>27</sup> (Tablo 2)

**TABLO 2:** Minimum içerik geçerlilik oranı değerleri.

Panelist Sayısı	Minimum Değer
5	0.99
6	0.99
7	0.99
8	0.78
9	0.75
10	0.62
11	0.59
12	0.56
13	0.54
14	0.51
15	0.49
20	0.42
25	0.37
30	0.33
35	0.31
40	0.29

Makalede Helsinki Deklerasyonu Prensipleri dikkate alınmıştır.

## BULGULAR

Lawshe (1975) çalışmasında, İGO ile ilgili kritik değerleri nasıl hesapladığına dair bir açıklamada bulunmamış bu değerlerin hesaplandığı temeli tanımlamamıştır. Farklı birinci tür hata değerleri için Binom dağılımına normal yaklaşıma bağlı olarak, panelist sayısının 5 ile 40 arasındaki değişen sayıları için kritik İGO değerleri Wilson ve diğ. (2012) tarafından hesaplanmıştır. Wilson ve diğ. (2012) tarafından İGO için kritik değer hesaplamalarında,

$$\dot{I}GO = \frac{n_{e,\alpha} - (N/2)}{N/2} = \frac{z_\alpha}{\sqrt{N}} \quad (2)$$

$$n_{e,\alpha} = z_\alpha * \frac{\sqrt{N}}{2} + \frac{N}{2} \quad (3)$$

eşitlikleri kullanılmıştır.<sup>25</sup> Eşitlik 2 ve Eşitlik 3'te kullanılan kritik  $z_\alpha$  değerleri Tablo 3'de verilmiştir. (Tablo 3).

Panelist sayısının 5 ile 40 arasında değişen değerleri için Binom dağılımına normal yaklaşıma dayanan İGO değerleri Tablo 4'te gösterilmiştir. (Tablo 4)

Tablo 5'te gösterilen Lawshe İGO için minimum panelist sayıları ise Eşitlik 3 ve Tablo 3'de verilen  $z_\alpha$  değerleri kullanılarak hesaplanmıştır. (Tablo 5)

**TABLO 3:**  $z_\alpha$  Değerleri.

Tek Yönlü Test $\alpha$	İki Yönlü Test $\alpha$	$z_\alpha$
0.10	0.20	1.281
0.05	0.10	1.645
0.025	0.05	1.959
0.01	0.02	2.299
0.005	0.01	2.590
0.001	0.002	3.090

**TABLO 4:** İçerik geçerlilik oranları için kritik değerler.

	<b>Tek Yönlü Test için Anlamlılık Düzeyi</b>					
	<i>0.10</i>	<i>0.05</i>	<i>0.025</i>	<i>0.01</i>	<i>0.005</i>	<i>0.001</i>
<i>N</i>	<b>İki Yönlü Test için Anlamlılık Düzeyi</b>					
	<i>0.20</i>	<i>0.10</i>	<i>0.05</i>	<i>0.02</i>	<i>0.01</i>	<i>0.002</i>
5	0.573	0.736	0.877	0.990*	0.990*	0.990*
6	0.523	0.672	0.800	0.950	0.990*	0.990*
7	0.485	0.622	0.741	0.879	0.974	0.990*
8	0.453	0.582	0.693	0.822	0.911	0.990*
9	0.427	0.548	0.653	0.775	0.859	0.990*
10	0.405	0.520	0.620	0.736	0.815	0.977
11	0.387	0.496	0.591	0.701	0.777	0.932
12	0.370	0.475	0.566	0.671	0.744	0.892
13	0.356	0.456	0.544	0.450	0.714	0.857
14	0.343	0.440	0.524	0.622	0.688	0.826
15	0.331	0.425	0.506	0.601	0.665	0.798
16	0.321	0.411	0.490	0.582	0.644	0.773
17	0.311	0.399	0.475	0.564	0.625	0.750
18	0.302	0.388	0.462	0.548	0.607	0.729
19	0.294	0.377	0.450	0.534	0.591	0.709
20	0.287	0.368	0.438	0.520	0.576	0.691
21	0.280	0.359	0.428	0.508	0.562	0.675
22	0.273	0.351	0.418	0.496	0.549	0.659
23	0.267	0.343	0.409	0.485	0.537	0.645
24	0.262	0.336	0.400	0.475	0.526	0.631
25	0.256	0.329	0.392	0.465	0.515	0.618
26	0.251	0.323	0.384	0.456	0.505	0.606
27	0.247	0.317	0.377	0.448	0.496	0.595
28	0.242	0.311	0.370	0.440	0.487	0.584
29	0.238	0.305	0.364	0.432	0.478	0.574
30	0.234	0.300	0.358	0.425	0.470	0.564
31	0.230	0.295	0.352	0.418	0.463	0.555
32	0.227	0.291	0.346	0.411	0.455	0.546
33	0.223	0.286	0.341	0.405	0.448	0.538
34	0.220	0.282	0.336	0.399	0.442	0.530
35	0.217	0.278	0.331	0.393	0.435	0.522
36	0.214	0.274	0.327	0.388	0.429	0.515
37	0.211	0.270	0.322	0.382	0.423	0.508
38	0.208	0.267	0.318	0.377	0.418	0.501
39	0.205	0.263	0.314	0.372	0.412	0.495
40	0.203	0.260	0.310	0.368	0.407	0.489

\*:  $IGO \geq 1$  olan değerler 0,99 olarak düzenlenmiştir.

**TABLO 5:** İçerik geçerlilik oranları için minimum panelist sayıları.

<i>N</i>	<b>Tek Yönlü Test için Anlamlılık Düzeyi</b>					
	<i>0.10</i>	<i>0.05</i>	<i>0.025</i>	<i>0.01</i>	<i>0.005</i>	<i>0.001</i>
<i>N</i>	<b>İki Yönlü Test için Anlamlılık Düzeyi</b>					
	<i>0.20</i>	<i>0.10</i>	<i>0.05</i>	<i>0.02</i>	<i>0.01</i>	<i>0.002</i>
5	4	4	5	5*	5*	5*
6	5	5	5	6	6*	6*
7	5	6	6	7	7	7*
8	6	6	7	7	8	8*
9	6	7	7	8	8	9*
10	7	8	8	9	9	10
11	8	8	9	9	10	11
12	8	9	9	10	10	11
13	9	9	10	11	11	12
14	9	10	11	11	12	13
15	10	11	11	12	13	13
16	11	11	12	13	13	14
17	11	12	13	13	14	15
18	12	12	13	14	14	16
19	12	13	14	15	15	16
20	13	14	14	15	16	17
21	13	14	15	16	16	18
22	14	15	16	16	17	18
23	15	15	16	17	18	19
24	15	16	17	18	18	20
25	16	17	17	18	19	20
26	16	17	18	19	20	21
27	17	18	19	19	20	22
28	17	18	19	20	21	22
29	18	19	20	21	21	23
30	19	20	20	21	22	23
31	19	20	21	22	23	24
32	20	21	22	23	23	25
33	20	21	22	23	24	25
34	21	22	23	24	25	26
35	21	22	23	24	25	27
36	22	23	24	25	26	27
37	22	24	24	25	26	28
38	23	24	25	26	27	29
39	23	25	26	27	28	29
40	24	25	26	27	28	30

\*:  $n_{e,a} \geq N$  olan değerler için  $n_{e,a} = N$  alınmıştır.

## TARTIŞMA

İçerik geçerliliği, “ölçeğin içeriği, değerlendirilen fenomen ile ilgili içeriğin evrenini temsil ediyor mu?” sorusu ile yönlendirilmektedir. İçerik geçerliliği, iki aşamalı (geliştirme ve karar niceleme) bir sürecin uygulanmasıyla, bir ölçeğin maddelerinin uygunluğunun belirlenmesidir. İçerik geçerliliğini belirlemek ve ölçmek için iki aşamalı bir işlem kullanmak, neredeyse tüm ölçeklerin doğrulanmasında esastır. Gelişim aşaması: bir ölçeğin geliştirilmesi öncelikle içerik geçerliliğinin değerlendirilmesi ile başlar. Geliştirme aşaması alan tanımlaması, madde üretimi ve ölçek oluşturma olarak ifade edilen üç adımdan oluşmaktadır. Karar-Niceleme Aşaması: içerik geçerliliğinin bu aşaması, hem bilişsel hem de duyuşsal ölçümler için aynı olan iki aşamaya sahiptir. Birinci aşamada ihtiyaç duyulan uzman sayısı belirlenir, ikinci aşamada ise içerik geçerliliği İGİ ve İGO değerleri hesaplanır.<sup>28</sup> İçerik geçerliliği, yargı ölçümü süreci ile değerlendirilir. Bir araştırmanın sonuçlarına güven duymak için, ölçeğin kusursuz bir şekilde uygulandığı zaman ölçmek istediğini tutarlı bir şekilde ölçmesi sağlanmalıdır. Kısacası, ölçek hem geçerli hem de güvenilir olmalıdır.<sup>12</sup> Bir ölçekte yer alacak maddelerin araştırma sonucunda verilecek kararları etkileyeceği açıktır. Dolayısıyla bir ölçekte yer alacak maddelerin belirlenmesinde uzmanlar paneli oluşturmak, ölçekte yer alacak maddelerin değerlendirilmesi amacıyla kullanılacak yöntemlerden biridir. Birçok araştırma konusunun karmaşık doğası nedeniyle, birçok disiplinin uzmanları ölçek geliştirmede içerik belirleyiciler olarak yararlı olabilir. Ancak içerik geçerliliği, uzmanların değerlendirmesine bağlı olduğundan uygun sayıda ve doğru uzmanları seçmek çok önemlidir.<sup>29</sup>

## SONUÇ

İçerik geçerliliğinin belirlenmesinde ihtiyaç duyulan uzman sayısını belirlemek, çoğunlukla keyfi olmaktadır. Uzmanların sayısı genellikle, herhangi bir prensibe bağlı olarak değil de ölçek geliştiricisinin ne kadar erişilebilir ve uygun insanı bulabileceğine bağlıdır. Her ne kadar bu uygulama yaygın olsa da, içerik geçerliliği belirlemesi için uzmanların seçiminde özel kurallar uygulanmalıdır. En az beş uzman yeterli olsa da bazı içerik alanlarında, bu kadar alan uzmanını bulmak ve işbirliğini sağlamak zor olabilir. Bu nedenle, en az üç uzman kullanılmalıdır. Kullanılabilecek asgari uzman sayısı belirlenmemiştir, ancak 10'u aşması muhtemel değildir.<sup>28</sup> İçerik geçerliliği için alternatif yöntemlerle (Cohen kappa katsayısı, Tinsley-Weiss T indeksi, James, Demaree ve Wolf indeksleri ve Lindell, Brandt ve Whitney indeksi) karşılaştırıldığında, Lawshe yöntemi basit ve kullanıcı dostudur. Yöntem yalnızca basit hesaplamalar gerektirmekte ve kritik değer belirlemek için araştırmacılara anlaşılması kolay bir tablo sunmaktadır.

### **Finansal Kaynak**

*Bu çalışma sırasında, yapılan araştırma konusu ile ilgili doğrudan bağlantısı bulunan herhangi bir ilaç firmasından, tıbbi alet, gereç ve malzeme sağlayan ve/veya üreten bir firma veya herhangi bir ticari firmadan, çalışmanın değerlendirme sürecinde, çalışma ile ilgili verilecek kararı olumsuz etkileyebilecek maddi ve/veya manevi herhangi bir destek alınmamıştır.*

### **Çıkar Çatışması**

*Yazarlar herhangi bir çıkar çatışması veya finansal destek bildirmemiştir.*

### **Yazar Katkıları**

**Fikir/Kavram:** İsmet Doğan, Nurhan Doğan; **Tasarım:** İsmet Doğan, Nurhan Doğan; **Denetleme/Danışmanlık:** İsmet Doğan, Nurhan Doğan; **Veri Toplama ve/veya İşleme:** İsmet Doğan, Nurhan Doğan; **Analiz ve/veya Yorum:** İsmet Doğan, Nurhan Doğan; **Kaynak Taraması:** İsmet Doğan, Nurhan Doğan; **Makalenin Yazımı:** İsmet Doğan, Nurhan Doğan; **Eleştirel İnceleme:** İsmet Doğan, Nurhan Doğan; **Kaynaklar ve Fon Sağlama:** İsmet Doğan, Nurhan Doğan; **Malzemeler:** İsmet Doğan, Nurhan Doğan

## KAYNAKLAR

1. Sirajudeen MS, Pillai PS, Shah UN, Mohan N. Content validity and inter-rater reliability of a checklist to assess the ergonomic practice of computer professionals. *Int J Rehabil Res.* 2012;1(2):11-8. [[Crossref](#)]
2. Ercaan İ, Kan İ. [Reliability and validity in the scales]. *Journal of the Uludağ University Medical Faculty.* 2004;30(3):211-6.
3. Bolarinwa OA. Principles and methods of validity and reliability testing of questionnaires used in social and health science researches. *Niger Postgrad Med J.* 2015;22(4):195-201. PMID: 26776330 [[Crossref](#)] [[PubMed](#)]
4. Zamanzadeh V, Ghahramanian A, Rassouli M, Abbaszadeh A, Alavi-Majid H, Nikanfar AR. Design and implementation content validity study: development of an instrument for measuring patient-centered communication. *J Caring Sci.* 2015;4(2):165-78. PMID: 26161370 [[Crossref](#)] [[PubMed](#)] [[PMC](#)]
5. Beck CT, Gable RK. Ensuring content validity: an illustration of the process. *J Nurs Meas.* 2001;9(2):201-15. PMID: 11696942 [[Crossref](#)] [[PubMed](#)]
6. Rubio DM, Berg-Weger M, Tebb SS, Lee ES, Rauch S. Objectifying content validity: conducting a content validity study in a social work research. *Soc Work.* 2003;27(2):94-104. [[Crossref](#)]
7. Baghestani AR, Ahmadi F, Tanha A, Meshkat M. Bayesian critical values for Lawshe's content validity ratio. *Meas Eval Couns Dev.* 2019;52(1):69-73. [[Crossref](#)]
8. Polit DF, Beck CT. The content validity index: are you sure you know what's being reported? Critique and recommendations. *Res Nurs Health.* 2006;29(5):489-97. PMID: 16977646 [[Crossref](#)] [[PubMed](#)]
9. Ozer İ, Fitzgerald SM, Sulbaran E, Garvey D. Reliability and content validity of an English as a foreign language (EFL) grade-level test for Turkish primary grade students. *Proc Soc Behv.* 2014;112:924-9. [[Crossref](#)]
10. Abdollahpour İ, Nedjat S, Noroozian M, Majdzadeh R. Performing content validation process in development of questionnaires. *Iran J Epid.* 2011;6(4):66-74.
11. Wynd CA, Schmidt B, Schaefer MA. Two quantitative approaches for estimating content validity. *West J Nurs Res.* 2003;25(5):508-18. PMID: 12955968 [[Crossref](#)] [[PubMed](#)]
12. Sireci SG. The construct of content validity. *Social Indic Res.* 1988;45(1-3):83-117. [[Crossref](#)]
13. Delegado-Rico E, Carretero-Dios H, Ruch W. Content validity evidences in test development: an applied perspective. *Int J Clin Hlth Psych.* 2012;12(3):449-60.
14. Ayre C, Scally AJ. Critical values for Lawshe's content validity ratio: revisiting the original methods of calculation. *Meas Eval Couns Dev.* 2014;47(1):79-86. [[Crossref](#)]
15. Polit DF, Beck CT, Owen SV. Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Res Nurs Health.* 2007;30(4):459-67. PMID: 17654487 [[Crossref](#)] [[PubMed](#)]
16. Yaghmaie F. Content validity and its estimation. *J Med Educ.* 2003;3(1):25-7.
17. Haynes SN, Richard DCS, Kubany ES. Content validity in psychological assessment: a functional approach to concepts and methods. *Psychol Assess.* 1995;7(3):238-47. [[Crossref](#)]
18. Koller İ, Levenson MR, Glück J. What do you think you are measuring? A mixed-methods procedure for assessing the content validity of test items and theory-based scaling. *Front Psychol.* 2017;8:126. PMID: 28270777 [[Crossref](#)]
19. Kassam-Adams N, Marsac ML, Kohser KL, Kenardy JA, March S, Winston FK. A new method for assessing content validity in model-based creation and iteration of eHealth interventions. *J Med Internet Res.* 2015;17(4):e95. PMID: 25881584 [[Crossref](#)] [[PubMed](#)] [[PMC](#)]
20. Gilbert GE, Prion S. Making sense of methods and measurement: Lawshe's content validity index. *Clin Simul Nurs.* 2016;12(12):530-1. [[Crossref](#)]
21. Ahari MN, Azad A, Alizadeh-Zarei M, Ebadi A, Parand A, Mohammadi P. Development and validity of the school interim competency of performance skill battery scale (SICPSBS). *Int J Pediatr.* 2018;6(11):8451-73.
22. Davies K, Bulsara MK, Ramelet AS, Monterosso L. Content validity testing of the ESAT: a decision aid tool for performing endotracheal suction in children. *Aust Crit Care.* 2018;31(1):23-30. PMID: 28254165 [[Crossref](#)] [[PubMed](#)]
23. Hadzaman NAH, Takim R, Nawawi AH, Mohamad Yusuwan N. Content validity of governing in Building Information Modelling (BIM) implementation assessment instrument. *IOP Conf Ser: Earth Environ Sci.* 2018;140(1): DOI: 10.1088/1755-1315/140/1/012105 [[Crossref](#)]
24. Halek M, Holle D, Bartholomeyczik S. Development and evaluation of the content validity, practicability and feasibility of the innovative dementia-oriented assessment system for challenging behaviour in residents with dementia. *BMC Health Serv Res.* 2017;17(1):554. PMID: 2880693 [[Crossref](#)] [[PubMed](#)] [[PMC](#)]
25. Wilson FR, Pan W, Schumsky DA. Recalculation of the critical values for Lawshe's content validity ratio. *Meas Eval Couns Dev.* 2012;45(3):197-210. [[Crossref](#)]
26. Shariati B, Keshavarz-Akhlaghi AA, Mohammadzadeh A, Seddigh R. The content validity of the cultural formulation interview (CFI). *Psychiatry J.* 2018;2018:3082823. PMID: 30631771 [[Crossref](#)] [[PubMed](#)] [[PMC](#)]
27. Lawshe CH. A quantitative approach to content validity. *Pers Psychol.* 1975;28(4):563-75. [[Crossref](#)]
28. Lynn MR. Determination and quantification of content validity. *Nurs Res.* 1986;35(6):382-5. PMID: 3640358 [[Crossref](#)]
29. Davis LL. Instrument review: getting the most from a panel of experts. *Appl Nurs Res.* 1992;5(4):194-7. [[Crossref](#)]